EEE 531: Semiconductor Device Theory I

Instructor: Dragica Vasileska

Department of Electrical Engineering
Arizona State University

Topics covered:
1. Introduction to MOSFET operation
1. Introduction

- First proposal for a MOSFET device was given by Lilienfeld and Heil in 1930.
- First operational device was made in 1947 at Bell Labs.
- Since then, the MOSFET dimensions have been continuously scaled to achieve more functions on a chip.

From SIA roadmap for semiconductors (1997)
- MOSFET is a four-terminal device. Basic device configuration is illustrated on the figures below.

**Side-view of the device**

*Image of side-view with labels: VDS, VGS, VBS, GATE, n-SOURCE, n-DRAIN, GATE OXIDE, INVERSION LAYER, p-SUBSTRATE, BACK CONTACT.*

**Top-view of the device**

*Image of top-view with labels: VDS, VGS, I_G = 0, I_D, n-SOURCE, n-DRAIN, W, L.*

**Basic device parameters:**
- channel length $L$
- channel width $W$
- oxide thickness $d_{ox}$
- junction depth $r_j$
- substrate doping $N_A$
There are basically four types of MOSFETs:

(a) $n$-channel, enhancement mode device

(b) $n$-channel, depletion mode device
(c) $p$-channel, enhancement mode device

(d) $p$-channel, depletion mode device
The role of the Gate electrode for \textit{n}-channel MOSFET:

\[ V_G = 0 \]

Positive gate voltage does two things:

1. Reduces the potential energy barrier seen by the electrons from the source and the drain regions.
2. Inverts the surface, and increases the conductivity of the channel.
The role of the **Drain** electrode for *n*-channel MOSFET:

1. **$V_G = 0, V_D > 0$**
   - Large potential barrier allows only few electrons to go from the source to the drain (subthreshold conduction)

2. **$V_G > V_T, V_D > 0$**
   - Smaller potential barrier allows a large number of electrons to go from the source to the drain
• Qualitative description of MOSFET operation:

(a) $V_G > V_T$, $V_D > 0$ (small)

Variation of electron density along the channel is small:

$$I_D \propto V_D$$

(b) $V_G > V_T$, $V_D > 0$ (larger)

Increase in the drain current reduces due to the reduced conductivity of the channel at the drain end.
(c) \( V_G > V_T, \ V_D = V_G - V_T \)

Pinch-off point. Electron density at the drain-end of the channel is identically zero.

(d) \( V_G > V_T, \ V_D > V_G - V_T \)

Post pinch-off characteristic. The excess drain voltage is dropped across the highly resistive pinch-off region denoted by \( \Delta L \).
• $IV$-characteristics (long-channel devices):

![Graph showing $IV$-characteristics with linear and saturation regions marked]
\[ N_A = 8 \times 10^{17} \text{ cm}^{-3}, \quad d_{ox} = 3 \text{ nm} \]
\[ V_G = 0.8 \text{ V}, \quad V_D = 20 \text{ mV}, \quad V_T = 0.33 \text{ V} \]
$V_G = 0.8 \text{ V, } V_D = 20 \text{ mV}$
\[ V_G = 0.8 \text{ V}, \; V_D = 0.2 \text{ V}, \; V_T = 0.33 \text{ V} \]
$V_G = 0.8 \, \text{V,} \quad V_D = 0.56 \, \text{V,} \quad V_T = 0.33 \, \text{V}$
$V_G = 0.8 \text{ V}, \ V_D = 0.56 \text{ V}, \ V_T = 0.33 \text{ V}$
\[ V_G = 0.8 \, \text{V}, \quad V_D = 0.9 \, \text{V}, \quad V_T = 0.33 \, \text{V} \]
$V_G = 0.8 \text{ V}, \ V_D = 1.56 \text{ V}, \ V_T = 0.33 \text{ V}$

**3D View**

**Contour plot**
$V_G = 0.8 \, \text{V}, \quad V_D = 1.56 \, \text{V}, \quad V_T = 0.33 \, \text{V}$
EEE 531: Semiconductor Device Theory I

Instructor: Dragica Vasileska

Department of Electrical Engineering
Arizona State University

Topics covered:

2. Gradual channel approximation for current calculation
   (A) square-law theory
   (B) bulk-charge theory
   (C) transconductance, output conductance and series resistance
   (D) limitations of the two models
2. Gradual channel approximation

- This model is due to Shockley.
- **Assumption:** The electric field variation in the direction parallel to the SC/oxide interface is much smaller than the electric field variation in the direction perpendicular to the interface:

\[ \left| \frac{dF_x}{dx} \right| \ll \left| \frac{dF_y}{dy} \right| \]
• Recall the expressions for the threshold voltage for real MOS capacitor:

\[
V_T = 2\phi_F + \frac{1}{C_{ox}} \sqrt{2qN_A k_s \varepsilon_0 (2\phi_F)} + V_{FB}
\]

Gate voltage:

Flat–band voltage:

\[
V_{FB} = \frac{1}{q} \phi_{MS} + \frac{Q_{it}}{C_{ox}} + \frac{Q_f}{C_{ox}} + \gamma_{ot} \frac{Q_{ot}}{C_{ox}} + \gamma_m \frac{Q_m}{C_{ox}}
\]

• Beyond the point that determines the onset of strong inversion (\(\phi_s = 2\phi_F\)), any excess charge on the gate balanced with excess charge in the semiconductor, is given by:

\[
Q_G = -(Q_B + Q_N) = C_{tot}(V_G - V_T) \rightarrow Q_N \approx -C_{ox}(V_G - V_T) - Q_B
\]

\[
Q_B = Q_B(\phi_s) - Q_B(2\phi_F)
\]

• Based on how we consider \(Q_B\), we have:

(A) Square-law theory: \(Q_B = 0\)

(B) Bulk-charge theory: \(Q_B \neq 0\)
(A) Square-law theory

- The charge on the gate is completely balanced by $Q_N(x)$, i.e:
  $$Q_N(x) \approx -C_{tot}[V_G - V_T - V(x)]$$

- Total current density in the channel:
  $$J_n = qn\mu_n F(x) + qD_n \frac{dn}{dx} \approx -qn\mu_n \frac{dV}{dx}$$

  Note: Total current density approximately equal to the electron current density (unipolar device).
• Integrating the current density, we obtain drain current $I_D$:

$$I_D = -\int_0^W dz \int_0^{y_c(x)} dy \left[ -qn(x, y)\mu_n(x, y)\frac{dV}{dx} \right]$$

$$= W \frac{dV}{dx} \int_0^{y_c(x)} qn(x, y)\mu_n(x, y)dy$$

$$\approx -Q_N(x)\mu_{eff} W \frac{dV}{dx}$$

$$\approx C_{ox} W \mu_{eff} [V_G - V_T - V(x)] \frac{dV}{dx}$$

Effective electron mobility, in which interface-roughness is taken into account.

High-resolution transmission electron micrograph of the interface between Si and SiO$_2$

• The role of interface-roughness on the low-field electron mobility:

![Graph showing the effect of doping and inversion charge density on mobility]

- Coulomb
- Phonon
- Interface-roughness

\[ N_A = 7 \times 10^{17} \text{ cm}^{-3} \]

Bulk samples

Si inversion layers
The experimentally observed universal mobility behavior is due to the dominant surface-roughness influence on the low-field electron mobility under strong inversion conditions.

Various models proposed for the variation of the effective transverse electric field upon the inversion ($N_s$) and depletion ($N_{depl}$) charge density:

- **Stern and Howard:** $F_{\text{eff}} \propto \left( \frac{11}{32} N_s + N_{depl} \right)$

- **Matsumoto and Uemura:** $F_{\text{eff}} \propto (0.5N_s + N_{depl})$

- **Krutsick and White:** $F_{\text{eff}} \propto \left[ 0.5N_s + \left( 1 - \frac{<Z>}{W} \right) N_{depl} \right]$

  (a=0.5 and b<1)
• There are several empirical expressions for the effective field dependence of the low-field electron mobility:

- **Effective-field dependence:**
  \[ \mu_{n,\text{eff}} = \frac{1105}{1 + \left(F_{\text{eff}} / 30.5\right)^{0.657}}, \quad \mu_{p,\text{eff}} = \frac{342}{1 + \left(F_{\text{eff}} / 15.4\right)^{0.617}} \]

- **Gate-voltage dependence:**
  \[ \mu_{n,\text{eff}} = \frac{\mu_0}{1 + \theta(V_{GS} - V_T)}, \quad 0.02 < \theta < 0.08 \quad [1/V] \]

• Universal mobility behavior:

![Diagram showing the relationship between \( N_{A1}, N_{A2}, N_{A3} \) and \( F_{\text{eff}}, \mu_{\text{eff}} \). The curves indicate \( N_{A3} > N_{A2} > N_{A1} \).]
• We now use the conservation of current argument, to get:

\[
\int_{0}^{L} I_D dx = W\mu_{\text{eff}}C_{ox} \int_{0}^{V_D} [V_G - V_T - V(x)] dV
\]

\[
\Rightarrow I_D = \frac{W\mu_{\text{eff}}C_{ox}}{L} \left[(V_G - V_T)V_D - \frac{1}{2}V_D^2\right], \text{ for } V_D \leq V_G - V_T
\]

• This last result represents the expression for the current, valid up to the pinch-off point.

• The current expression beyond the pinch-off point, at which

\[
V_D = V_G - V_T \rightarrow Q_N(L) = 0
\]

is obtained by construction, to give:

\[
I_D = \frac{W\mu_{\text{eff}}C_{ox}}{2L} (V_G - V_T)^2, \text{ for } V_D \geq V_G - V_T
\]
(B) Bulk-charge theory

- Square-law theory assumes that excess $Q_G$ is solely balanced by $Q_N$, and that $W=W_T=\text{const.}$ along the channel.
- The violation of this assumption is clearly shown on the figure below:

![Graph showing the violation of assumption with source and drain labels and dimensions in microns.]

- $N_A = 8 \times 10^{17} \text{cm}^{-3}$, $d_{ox} = 3 \text{nm}$
- $V_G = 0.8 \text{V}$, $V_D = 1.56 \text{V}$, $V_T = 0.33 \text{V}$
• Taking into account the contribution by the bulk charges, the more exact electron density is given by:

\[ Q_N(x) = -C_{ox}[V_G - V_T - V(x)] - Q_B(x) \]

where:

\[ Q_B(x) = -qN_A[W(x) - W_T] \]

\[ W(x) = \sqrt{\frac{2k_s\varepsilon_0}{qN_A}[2\phi_F + V(x)]}, \quad W_T = \sqrt{\frac{2k_s\varepsilon_0}{qN_A}(2\phi_F)} \]

• Following the same steps as in the square-law theory, we get:

\[ I_D = \frac{W_{\mu_{\text{eff}}}C_{ox}}{L} \left\{ \left[ (V_G - V_T)V_D - \frac{1}{2}V_D^2 \right] - \frac{4}{3}V_W\phi_F \left[ \left( 1 + \frac{V_D}{2\phi_F} \right)^{3/2} - \left( 1 + \frac{3V_D}{4\phi_F} \right) \right] \right\} \]

\[ V_W = \frac{\sqrt{2k_s\varepsilon_0qN_A(2\phi_F)}}{C_{ox}} = \frac{W_TqN_A}{C_{ox}} \]
The pinch-off voltage $V_{D_{sat}}$, is obtained from the condition that $Q_N(L) = 0$, which gives:

$$
V_{D_{sat}} = V_G - V_T - V_W \left\{ \sqrt{\frac{V_G - V_T}{2\phi_F}} + \left(1 + \frac{V_W}{4\phi_F}\right)^2 - \left(1 + \frac{V_W}{4\phi_F}\right) \right\}
$$

Comparison between the two theories:
(C) Transconductance, drain conductance, series resistance

• Using square-law theory and for small $V_D$, we get:

$$I_D \approx \frac{W \mu_{\text{eff}} C_{ox}}{L} \left( V_G - V_T \right) V_D$$

transconductance:  \( g_m = \frac{\partial I_D}{\partial V_G} \bigg|_{V_D=\text{const.}} \approx \frac{W \mu_{\text{eff}} C_{ox}}{L} V_D \)

drain conductance:  \( g_d = \frac{\partial I_D}{\partial V_D} \bigg|_{V_G=\text{const.}} \approx \frac{W \mu_{\text{eff}} C_{ox}}{L} \left( V_G - V_T \right) \)

• For large values for $V_D$, we get:

$$g_m = \frac{W \mu_{\text{eff}} C_{ox}}{L} \left( V_G - V_T \right), \quad g_d = 0$$
• The simplified low-frequency and high-frequency MOSFET small-signal equivalent circuits are shown below:

![Low-frequency MOSFET equivalent circuit](image1)

![High-frequency MOSFET equivalent circuit](image2)

- Low-frequency
- High-frequency

• The cut-off frequency (frequency for which the short-circuit current gain equals one) is given by:

\[
f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})} \rightarrow \frac{1}{2\pi t_{tr}}, \quad t_{tr} = \frac{L}{v_{eff}}
\]

Transit time of the electrons in the channel
• Assumption made in the previous derivations is that the entire voltage drop is across the channel.

• In real devices, both the drain and the source resistances $R_d$ and $R_s$ may play an important role, thus limiting the device performance.

\[
V_{GS} = V_G + R_s I_D \\
V_{DS} = V_D + (R_s + R_d) I_D
\]
• These series resistances modify the transconductance and the output conductance:

$$g_m = \frac{\partial I_D}{\partial V_G} \bigg|_{V_D=\text{const.}} = \frac{g_{m0}}{1 + g_{m0}R_s + g_{d0}(R_s + R_d)}$$

$$g_d = \frac{\partial I_D}{\partial V_D} \bigg|_{V_G=\text{const.}} = \frac{g_{d0}}{1 + g_{m0}R_s + g_{d0}(R_s + R_d)}$$

**Transfer characteristic**

**Output characteristic**
(D) Limitations of the square-law and bulk-charge theories:

(1) They do not include the subthreshold region.

(2) Both theories do not self-saturate. One must obtain the post-pinch off characteristics by construction.

(3) The exact charge model self-saturates and naturally includes the subthreshold.
EEE 531: Semiconductor Device Theory I

Instructor: Dragica Vasileska

Department of Electrical Engineering
Arizona State University

Topics covered:

3. MOSFET subthreshold characteristics
3. Subthreshold characteristics

- **Above threshold** → current is governed by the channel resistance. Drift current dominates.
- **Below threshold** → current is barrier limited. Diffusion current dominates.

- Recall PN junctions: Current depends exponentially on the applied bias → the excess minority carrier density at the edge of the SCR depends exponentially on the applied voltage.

- MOSFETs appear to have the same problem → electrons injected into the channel from the source also become minority carriers and need to climb a potential barrier. The important point in MOSFET operation is that the gate electrode eliminates this barrier. Only in subthreshold the barrier will play significant role in limiting the current flow in the device.
Accumulation condition

Depletion condition:
- small inversion charge
- small drain current
- slight potential gradient from source to drain

Inversion condition
• Subthreshold current flows when the device is in weak inversion:
  \[ \phi_F < \phi_s < 2\phi_F \]

• In long-channel devices, the voltage drop \( V_D \) is entirely across the drain-substrate junction, which makes the in-plane component of the field \( (F_x) \) small, and current is diffusion limited:
  \[ I_D \approx -qA_{\text{eff}}D_n \frac{dn}{dx} \]

  Effective cross-section for the subthreshold current

• If the electron diffusion length is much larger than \( L \), the electron density varies linearly when going from the source \( (n_{ss}) \) to the drain \( (n_{sd}) \):
  \[ n(x) \approx n_{ss} - (n_{ss} - n_{sd}) \frac{x}{L} \]

  \[ n_{ss} = n(0) \rightarrow \text{source end} \]
  \[ n_{sd} = n(L) \rightarrow \text{drain end of the channel} \]
• Electron densities at the source and drain end of the channel:

Source end of the channel:

\[ n_{ss}(y) = n_i \exp\left[ \frac{E_F - E_i(y)}{k_B T} \right] \]

\[ = n_{po} \exp\left[ \frac{\varphi(y)}{V_T} \right] \]

Drain end of the channel:

\[ n_{sd}(y) = n_i \exp\left[ \frac{E_{Fn} - E_i(y)}{k_B T} \right] \]

\[ = n_{po} \exp\left[ \frac{\varphi(y) - V_D}{V_T} \right] \]
• We now interpolate the potential variation along the depth with:

\[
\varphi(y) = \varphi_s - F(y = 0)y = \varphi_s - F_s y, \quad F_s = \frac{1}{k_s \varepsilon_0} \sqrt{2qN_A k_s \varepsilon_0 \varphi_s}
\]

Surface potential

Surface electric field

• Consider now the electron density at the source end of the channel:

\[
n_{ss}(y) = n_{po} \exp\left[\frac{\varphi_s - F_s y}{V_T}\right] = n_{po} \exp\left[\frac{\varphi_s}{V_T}\right] n_{po} \exp\left[-\frac{F_s y}{V_T}\right]
\]

This term suggests that the effective thickness of the inversion layer along the depth (\(y\)-direction) is:

\[
y_{\text{eff}} = \frac{V_T}{F_s} = \frac{k_s \varepsilon_0 V_T}{qN_A} \sqrt{\frac{qN_A}{2k_s \varepsilon_0 \varphi_s}}
\]
• Substituting these results into the diffusion current expression gives:

\[ I_D = \mu_n k_s \varepsilon_0 \left( \frac{W}{L} \right) V_T^2 \left( \frac{n_i}{N_A} \right)^2 \sqrt{\frac{V_T}{\varphi_s \sqrt{2 L_D p}}} e^{\varphi_s / V_T} \left( 1 - e^{-V_D / V_T} \right) \]

Important notes:
- The subthreshold current is nearly independent of \( V_D \) if \( V_D > 3 V_T \)
- The subthreshold current depends exponentially on \( \varphi_s \)
- The subthreshold current depends upon the ratio \( W/L \)

• For \( V_D > 3 V_T \), the subthreshold current simplifies to:

\[ I_D = b \sqrt{\frac{V_T}{\varphi_s}} e^{\varphi_s / V_T} \rightarrow \ln(I_D) = \ln(b) + \frac{1}{2} \ln \left( \frac{V_T}{\varphi_s} \right) + \frac{\varphi_s}{V_T} \]

\[ d \ln(I_D) = \frac{1}{V_T} \left[ 1 - \frac{V_T}{2 \varphi_s} \right] d\varphi_s \]
To obtain the expression for the subthreshold swing $S$, we now utilize the relationship between the surface potential and the gate voltage:

$$V_G = \phi_s - \frac{Q_s(\phi_s)}{C_{ox}} + V_{FB} = \phi_s + \frac{1}{C_{ox}} \sqrt{2qN_A k_s \varepsilon_0 \phi_s} + V_{FB}$$

$$\frac{dV_G}{d\phi_s} = 1 + \frac{C_s(\phi_s)}{C_{ox}}$$

Combining the results for $d\ln(I_D)$ and $dV_G$, gives:

$$S = \frac{dV_G}{d \log(I_D)} = V_T \ln 10 \left[ 1 + \frac{C_s(\phi_s)}{C_{ox}} \right] \left[ 1 - \frac{2}{a^2} \left( \frac{C_s}{C_{ox}} \right)^2 \right]^{-1}$$

where $a = \sqrt{2k_s \varepsilon_0 / (L_{Dp} C_{ox})}$. If $a >> C_s/C_{ox}$ then:

$$S = \frac{dV_G}{d \log(I_D)} = V_T \ln 10 \left[ 1 + \frac{C_s(\phi_s)}{C_{ox}} \right] \rightarrow S_{\text{min}} = 60 \text{ mV / decade}$$
• The subthreshold swing tells us how fast we can turn the device off. Devices with good turn-off characteristics have subthreshold swings between 70 and 80 $mV/\text{decade}$.
EEE 531: Semiconductor Device Theory I

Instructor: Dragica Vasileska

Department of Electrical Engineering
Arizona State University

Topics covered:

4. Threshold voltage adjustment
4. Threshold voltage adjustment

- The threshold voltage is an important device parameter whose values should not fall outside certain prescribed limits.

- The conflicting design requirements between reducing $V_D$, $V_T$, $I_{leak}$, $C$, and increasing performance are schematically shown in the figure below:

  ![Diagram](image)

  - Reduce $P_{\text{switch}} \sim V_{DD}^2$
  - Reduce $P_{\text{leak}} \sim V_T$
  - Increase Performance
  - Design space is shrinking!
We now recall the expressions for the threshold voltage to understand which parameters are easily and reproducibly changed to give the desired threshold voltage:

\[
V_T = 2\phi_F - \frac{Q_s (2\phi_F)}{C_{ox}} + V_{FB} = 2\phi_F + \frac{1}{C_{ox}} \sqrt{2qN_A k_s \epsilon_0 (2\phi_F)} + V_{FB}
\]

\[
V_{FB} = \frac{1}{q} \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{Q_{it} (2\phi_F)}{C_{ox}} - \gamma \frac{Q_m}{C_{ox}}
\]

The variation of \(N_A\) makes significant contribution here.

\[
\Phi_{MS} = \Phi_M - \left[ \chi_{sc} + (E_C - E_F)_{bulk} \right]
\]

\[
= \Phi_M - \chi_{sc} - 0.5E_g - k_B T \ln (N_A / n_i)
\]

For each factor of 10 in doping concentration change, this term changes by 2.3 \(k_B T\) (not very much).
• To summarize, threshold voltage controlling parameters are:
  (A) Substrate doping
  (B) Substrate bias via $\phi_s$
  (C) Oxide thickness (useful for major threshold voltage control, but not for threshold voltage adjustments).

(A) Substrate doping

• The key process parameter for threshold voltage control is the substrate doping.

• For general, non-uniform doping density, the relationship between $V_G$ and $\phi_s$ is:

$$V_G = \phi_s - \frac{Q_s(\phi_s)}{C_{ox}} + V_{FB} = \phi_s + \frac{q}{C_{ox}} \int_0^W N_B(y)dy + V_{FB}$$

Depletion region depth

Acceptee-type doping density
• The substrate doping concentration can be modified, for example, using ion implantation process.

• Two limiting cases are interesting to consider:
  - very shallow heavily-doped surface layer
  - general ion-implanted impurity profile

• For very shallow surface layer, we have:

\[ N_B(y) = N_A + D_i \delta(0) \]

\[
V_G = \varphi_s + \frac{q}{C_{ox}} \int_0^W [N_A + D_i \delta(0)] dy + V_{FB}
\]

\[
= \varphi_s + \frac{qN_A W}{C_{ox}} + \frac{qD_i}{C_{ox}} + V_{FB}
\]

Dose: # of atoms per unit area

Acceptors → positive shift

Donors → negative shift
• For ion-implanted impurity profiles:

\[ N_i(y) = \frac{D_i}{\sqrt{2\pi}\Delta R_p} \exp \left[ -\frac{(y - R_p)^2}{2\Delta R_p^2} \right] \]

- \( D_i \rightarrow \) ion dose
- \( R_p \rightarrow \) range
- \( \Delta R_p \rightarrow \) straggle

\[ D_i = (N_i - N_A)d_i \]

Real ion-implanted profile

Approximation to the real profile
• Two special cases need to be considered for the step doping profile:

\[ d_i < W \rightarrow \text{all implanted ions are in the SCR} \]

\[ V_T = V_{FB} + 2\phi_F + \frac{qD_i}{C_{ox}} + \frac{1}{C_{ox}} \sqrt{2qN_A k_s \varepsilon_0 \left(2\phi_F - \frac{qD_id_i}{2k_s \varepsilon_0}\right)} \]

\[ d_i > W \rightarrow \text{the depth over which we have implanted ions exceeds the SCR depth} \]

\[ V_T = V_{FB} + 2\phi_F + \frac{1}{C_{ox}} \sqrt{2qN_i k_s \varepsilon_0 (2\phi_F)} \]

Need to substitute for different doping
(B) Substrate bias

- Reverse, or back-biasing, is another method that has been employed to adjust the threshold voltage.

\[ \phi_s = 2\phi_F \rightarrow \text{onset of strong inversion} \]

\[ \phi_s = 2\phi_F - V_{BS} \rightarrow \text{onset of strong inversion} \]

Energy-band diagram for \( V_{BS} = 0 \)

Energy-band diagram for \( V_{BS} \neq 0 \)
• From the energy-band diagrams shown in the previous slide, it is clear that the surface will invert when $\phi_s = 2\phi_F - V_{BS}$

• The threshold voltage is then given by:

$$V_T = V_{FB} + (2\phi_F - V_{BS}) + \frac{1}{C_{ox}} \sqrt{2qN_i k_s \varepsilon_0 (2\phi_F - V_{BS})}$$

Important notes:

➔ Back-biasing always increases $V_T$

➔ Current-voltage relations remain the same provided:

$$2\phi_F \rightarrow 2\phi_F - V_{BS}$$

• It is generally desirable to have low substrate-bias sensitivity (shallow channel implant + appropriate ion dose)
(C) Threshold voltage extraction

- Criterion 1:
  \[ V_T = V_G \text{ for which } I_D = 10 \, \mu A \] (Not accurate, but easy to use)

- Criterion 2:

  The intercept on the \( I_D = 0 \) axis gives the threshold voltage.
EEE 531: Semiconductor Device Theory I

Instructor: Dragica Vasileska

Department of Electrical Engineering
Arizona State University

Topics covered:

5. Small-geometry effects
   - Subthreshold slope increase
   - Short-channel effects
   - Narrow-width effects
   - Hot-carrier effects
   - Discrete impurity effects
   - Schematic description of realistic device structure
5. Short-Channel Effects in Scaled Si-MOSFETs

- Increase in the subthreshold current:
  - drain-induced barrier lowering

- Threshold voltage modification:
  - short-channel effects
  - narrow-width effects
  - quantum-mechanical charge description

- Transconductance degradation:
  - finite inversion layer capacitance
  - depletion of the polysilicon gates

- Parasitic BJT action:
  - punch-through effect
  - substrate current increase due to carrier multiplication and regenerative feedback

- Hot-carrier effects:
  - oxide charging
  - velocity saturation
  - velocity overshoot
  - ballistic transport

- Classical statistical effects:
  - random dopant fluctuations

Minimum channel-length below which significant short-channel effects are expected to occur:

\[ L_{\text{min}} = A \left[ \frac{r_J T_{\text{ox}} (W_S + W_D)^2}{\kappa N_A} \right]^{1/3} \]
(A) Subthreshold Current and Threshold Voltage

- For a long-channel device, the subthreshold current is independent of $V_D$, once $V_D > 3V_T$
- The situation is rather different for small devices, where one observes:
  - increase in the subthreshold swing
  - drain voltage dependence of the subthreshold current due to Drain Induced Barrier Lowering (DIBL)
(B) Short-channel effects

- In all our previous analysis, it was assumed that the gate charge $Q_G$ equals the sum of the electron charge $Q_N$ and depletion charge $Q_B$.

- Even in long-channel device, this is not strictly true. Part of the channel charge is controlled by the source and drain, not by the gate.

Only the charge in the shaded area is controlled by the gate.
• With increasing $V_D$, the amount of charge controlled by the drain increases → lower gate voltage is needed to invert the channel, i.e. $V_T$ decreases.

• This, in turn, leads to increase in the drain current.
• Graphical description of the problem:

\[ r_j + y = \sqrt{(r_j + W)^2 - W^2} = r_j \sqrt{1 + 2W / r_j} \]
\[ L' = L - 2y = L - 2r_j \left[ \sqrt{1 + 2W / r_j} - 1 \right] \]

• Total charge controlled by the gate (gate-width Z) is:

- long-channel device: \( Q_{BZL} = -qN_A ZL \)
- short-channel device:

\[ Q'_{BZL} = -qN_A Z \frac{1}{2} (L + L') \rightarrow Q'_B = -qN_A \frac{L + L'}{2L} \]
- Recall the expression for the threshold voltage:

\[ V_T = 2\phi_F - \frac{Q_B}{C_{ox}} + V_{FB} \]

- The threshold voltage shift is then given by:

\[
\Delta V_T = -\frac{Q'_B - Q_B}{C_{ox}} = -qN_AW \cdot \frac{r_j}{L} \left[ \sqrt{1 + \frac{2W}{r_j}} - 1 \right] = -V_W \frac{r_j}{L} \left[ \sqrt{1 + \frac{2W}{r_j}} - 1 \right]
\]
(C) Narrow-width effects

- The channel width also affects the threshold voltage, due to the additional lateral component in the CSR width (the gate controlled region extends on the sides, which gives:

\[ L Z Q'_B = L \left[ Q_B Z - 2qN_A \frac{1}{4} \pi W_T^2 \right] \]

\[ = L \left[ Q_B Z - \frac{1}{2} qN_A \pi W_T^2 \right] \]

- The difference in charge is:

\[ Q'_B - Q_B = -\frac{qN_A \pi W_T^2}{2Z} \]

- This gives the following threshold voltage shift:

\[ \Delta V_T = -\frac{Q'_B - Q_B}{C_{ox}} = \frac{qN_A \pi W_T^2}{2ZC_{ox}} = V_W \cdot \frac{\pi W_T}{2Z} \]
• Graphical representation of the $\Delta V_T$ shift:

![Graphical representation of $\Delta V_T$ shift](image)

Variation of $V_T$ with channel width $Z$

• The combined effect of both short-channel and narrow-width effect gives:

$$\Delta V_T = V_W \left\{ \xi \frac{W}{Z} - \left[ \sqrt{1 + 2 \frac{W}{r_j}} - 1 \right] \left[ 1 + \xi \frac{W}{Z} \right] \frac{r_j}{L} \right\}$$

For more detailed expression and actual doping profiles, numerical analysis is needed.
EEE 531: Semiconductor Device Theory I

Instructor: Dragica Vasileska

Department of Electrical Engineering
Arizona State University

**Topics covered:**

5. Small-geometry effects
   - Subthreshold slope increase
   - Short-channel effects
   - Narrow-width effects
   - Hot-carrier effects
   - Discrete impurity effects
   - Schematic description of realistic device structure
(D) Hot-carrier effects

• Under high bias conditions, the electrons at the drain end of the channel become very energetic (hot). This can give rise to several undesirable effects, such as:
  
  ★ Velocity saturation
  ★ Punch-through effect
  ★ Snapback breakdown (parasitic BJT action)
  ★ Oxide charging and tunneling currents
  ★ Velocity overshoot effect

• The ballistic transport effects, such as velocity overshoot lead to smaller transit time of the carriers and are, therefore, desirable effects.
Velocity saturation

- **Long-channel devices** → $I_D-V_D$ curves nearly constant in saturation.
- **Short-channel device** → Electric fields become very high and the drift-velocity becomes constant (mobility decreases).

Drift velocity and average electron energies for bulk Si
• A simplified expression is obtained using:

\[ I_D = -qA_{\text{eff}} n v_d \]

\[ A_{\text{eff}} = Z_y^\text{eff} \]

i.e. the velocity-limited drain current equals to:

\[ I_D = -qZy_{\text{eff}} n v_d = -qy_{\text{eff}} n Z v_d = Z v_d C_{\text{ox}} (V_G - V_T) \]

\[ \frac{Q}{Q_N} \]

• Comparing the above expression with the mobility-limited one for a long-channel device, we get:

\[ v_{\text{sat}} \leftrightarrow \frac{\mu_{\text{eff}} (V_G - V_T)}{2L} \]

For \( V_G - V_T = 5 \) V we get \( L = 1.25 \) \( \mu \)m

(\( v_{\text{sat}} = 10^7 \) cm/s)
Channel length = 0.2 \mu m

Experimental device

Calculated IV-characteristics with velocity saturation

Calculated IV-characteristics without velocity saturation effect
**Punch-through effect**

- Occurs when the source and drain depletion regions touch.
- The majority electrons in the source get injected into the depletion region where they are swept by the high electric field.
- Drain current is dominated by the space-charge current ($\sim V_D^2$) and not by the inversion layer current.
- To eliminate this effect, a punch-through stop is used. This is done with deep ion-implantation process.
Marginal long-channel device (Device 1)

Device with severe short-channel effects (Device 2)
Snapback breakdown

- Electrons near the drain region impact ionize, i.e. generate electron-hole pairs.
- Electrons are swept by the drain, and holes go to the substrate, forward-biasing the source-substrate junction. This leads to higher electron injection into the substrate.
- More electrons in the substrate means more impact ionization, i.e. positive feedback effect.
- The snapback portion comes because the source-substrate-drain form a BJT in parallel to the MOSFET that exhibits negative resistance or snapback.
Oxide charging and tunneling currents

- Oxide charging, or charge injection and trapping, is another undesirable effect.
- Electrons at the drain end of the channel have sufficient energy to overcome the barrier at the \( Si/SiO_2 \) interface and be trapped in the oxide.

Since the effect is cumulative, it limits the useful ‘life’ of the device. LDD regions are used to reduce oxide charging.
- Tunneling currents lead to gate leakage. The three types of tunneling processes are schematically shown below:

- For $t_{ox} \geq 40$ Å, Fowler-Nordheim (FN) tunneling dominates.
- For $t_{ox} < 40$ Å, direct tunneling becomes important.
- $I_{dir} > I_{FN}$ at a given $V_{ox}$ when direct tunneling active.
- For given electric field:
  - $I_{FN}$ independent of oxide thickness.
  - $I_{dir}$ dependent on oxide thickness.

$V_{ox} > \phi_B$ for FN tunneling.
$V_{ox} = \phi_B$ for FN/Direct tunneling.
$V_{ox} < \phi_B$ for Direct tunneling.
As oxide thickness decreases, gate current becomes more important. It eventually dominates the off-state leakage current ($I_D$ at $V_G = 0$).
**Velocity overshoot effect**

- We can describe the motion of the electrons between collisions by simple Newton’s Law:
  \[ m \ast \frac{dv}{dx} = -qF_x \]

- In a simplified approach (momentum balance equation for an average carrier) that neglects diffusion, we have:
  \[ m \ast \frac{dv}{dx} = -qF_x - \frac{m \ast v}{\tau_m} \]

- For a uniform electric field applied at \( t=0 \), the solution of the above equation is of the form:
  \[ v_{dx}(t) = \frac{q \tau_m}{m \ast F_x} \left[ e^{-t/\tau_m} - 1 \right] \]
• When steady-state has been reached, the electrons have traveled the distance:

\[ d = \tau_m \int_0^\tau_m v_{dx}(t) dt = -\frac{q\tau_m^2}{em} F_x \]

• For \( F_x = 10 \text{ kV/cm} \), we have that:
  
  \( \rightarrow d \approx 200\text{Å} \) for electrons in Si

• Why do we observe velocity overshoot?
  
  (1) The energy relaxation time is larger than momentum relaxation time

  (2) At first, the electric field simply displaces the distribution function with little change on its shape.

  (3) Later on, collisions broaden the distribution, the electron temperature increases and drift velocity drops.

• Velocity overshoot effect reduces the electron transit time, i.e. leads to faster devices.
EEE 531: Semiconductor Device Theory I

Instructor: Dragica Vasileska

Department of Electrical Engineering
Arizona State University

Topics covered:

5. Small-geometry effects
   - Subthreshold slope increase
   - Short-channel effects
   - Narrow-width effects
   - Hot-carrier effects
   - Discrete impurity effects
   - Schematic description of realistic device structure
(E) Discrete impurity effects

- In ultra-small devices, discrete impurity effects become important. This is clearly seen in the experimental measurements shown below.

Experimental investigations:

• The atomistic nature of the impurity atoms leads to fluctuations in the potential.

• The potential fluctuations affect the magnitude of the current and the threshold voltage for devices fabricated on a same chip.

Simulated devices have:
L = 0.1 μm, Z = 50 nm
The applied voltages are:
V_G = 0 V, V_D = 10 mV

Conduction band edge along depth and parallel to the SC/oxide interface
• Influence on the subthreshold transfer characteristics:

The spread of the transfer characteristics along the gate axis is due to the non uniformity of the potential barrier that allows for early turn-on at some parts of the channel.
Threshold voltage fluctuations are clearly seen on the example shown below, where we plot the # of channel dopant atoms (~I_D) as a function of V_G for two devices taken at the ends of the distribution of the statistical ensemble of 30 devices considered in this study. All devices considered here have identical geometry.
• Scatter plots of the threshold voltage versus the number of dopant atoms clearly show that devices with larger channel width have smaller threshold voltage fluctuations. We use $L_G=50$ nm, $N_A=5 \times 10^{18}$ cm$^{-3}$, $T_{ox}=2$ nm in this study.

• Analytical expression for the threshold voltage standard deviation:

$$\sigma_{V_{th}} \approx \frac{4}{\sqrt{3}} \left[ \frac{q^3 \epsilon_{Si} \Phi_F}{\sqrt{4q \epsilon_{Si} \Phi_F N_A}} + \frac{T_{ox}}{\epsilon_{ox}} \right] \frac{4 \sqrt{N_A}}{L_{eff} W_{eff}}$$
• Variation of the threshold voltage standard deviation with substrate doping, oxide thickness and device width is shown below.

**Approach 1 [1]:**

\[
\sigma_{Vth} = \frac{4\sqrt{3} \varepsilon_{Si} \Phi_B}{\sqrt{2}} T_{ox} \frac{4\sqrt{N_A}}{\varepsilon_{ox} \sqrt{L_{eff} W_{eff}}} ; \quad \phi_B = \frac{k_B T}{q} \ln \left( \frac{N_A}{n_i} \right)
\]

**Approach 2 [2]:**

\[
\sigma_{Vth} \approx \frac{4\sqrt{3} \varepsilon_{Si} \Phi_B}{\sqrt{3}} \left[ \frac{k_B T / q}{\sqrt{4q\varepsilon_{Si} \phi_B N_A}} + \frac{T_{ox}}{\varepsilon_{ox}} \right] \frac{4\sqrt{N_A}}{\sqrt{L_{eff} W_{eff}}}
\]


To understand the role that the **position** of the impurity atoms plays on the threshold voltage fluctuations, statistical ensembles of 5 devices from the **low-end, center** and the **high-end** of the distribution were considered.

Significant **correlation** was observed between the threshold voltage and the number of atoms that fall within the first **15 nm depth** of the channel.
- Impurity distribution in the channel also affects the carrier mobility and saturation current of the device.
- Significant correlation was observed between the drift velocity (saturation current) and the number of atoms that fall within the first 10 nm depth of the channel.
Below shown is a schematic of realistic device structures and highlight of some critical issues in device fabrication.

- Low resistivity
- Low K dielectric
- Ti, Co silicide?
- Gate depletion
- Oxide integrity
- Oxide thickness
- Short channel effects (e.g., DIBL)
- Epi wafers?
• Cross-sectional micrographs of a 60-nm MOSFET built at Bell Labs with 1.2 nm gate oxide.

In production 2010:
• 64-Gb DRAM
• 200-GHz transistor speeds
• 10-GHz processor clocks